Statistical Tools For Developing Nutrient Criteria

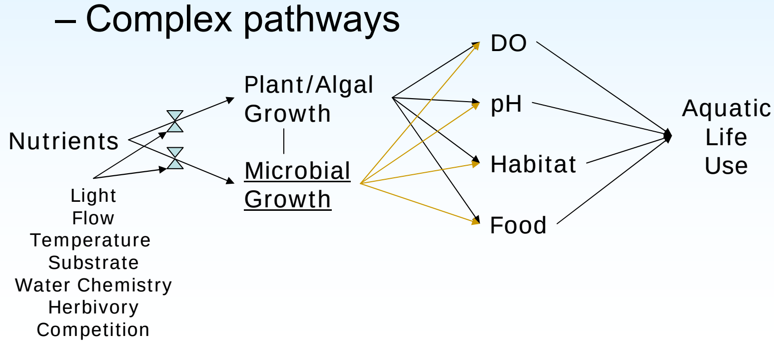
Michael J. Paul
Tetra Tech, Inc.

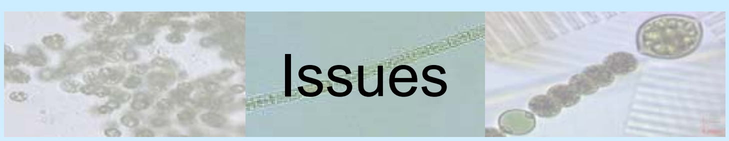
Nutrient Criteria Considerations

- Goal:
  - restore/protect uses
  - develop protective and defensible criteria
    - transparency and reproducibility

Issues

- Toxic criteria models hard to apply
- Field data
  - Multiple stressor issue
  - Complex pathways





Issues

- Given complexity, single analytical approach is hard

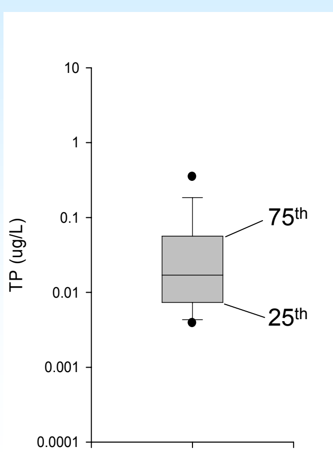
- Alternative – use multiple approaches and select criterion based on weighting these different values

Approaches

- Multiple Lines of Evidence
  - – Frequency distribution approaches
  - – Modeled and inferred reference expectation
  - – Nutrient-response models
  - – Mechanistic/Other Models
  - – Other Studies

USEPA 2000

Frequency Distribution



- Reference
- All
- Attaining

Distributional Statistics

- Central Tendency
  - Mean/Average = sum of all values/number of values
  - Median = the 50th percentile of values ranked in order
- Spread
  - Variance = average squared deviation of values from the mean
  - Standard deviation = square root of the variance

Distributional Statistics

- Spread
  - Percentile = fraction of values ranked in order
  - Quartile = 25th, 50th, and 75th percentiles
  - Inter-quartile range = range between 25th and 75th quartiles
  - Outlier = value numerically distant from rest

Distributional Statistics

- Mean
- Median
- Quartiles
- Standard deviation

1. 4
2. 6
3. 8
4. 10
5. 10
6. 11
7. 11
8. 12
9. 12
10. 12
11. 13
12. 13
13. 14
14. 14
15. 16
16. 18
17. 20

Mean = (4+6+…+17)/17 = 12

25th Percentile = 10

Median = 12

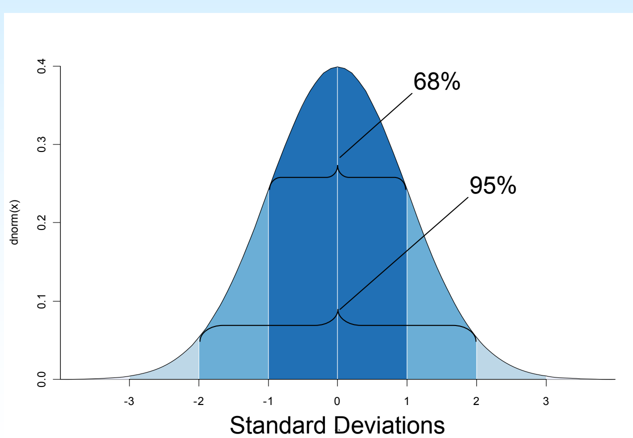
75th Percentile = 14

Variance = {(4-12)+(6-12)+…+(20-12)}/17

Standard Deviation = $\sqrt{\text{Variance}}$

Distributional Statistics

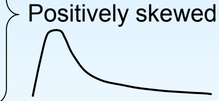
- Normal Distribution
- Many statistics assume normally distributed data
- Kolmogorov-Smirnov test
- Shapiro-Wilk test



Transformations

- Commonly used
  - 1/x
  - Common Log
  - Natural Log
  - Square-root
  - Fourth-root

  Positively skewed

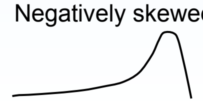
  

  - Arc-sine square-root

  Percentile Data

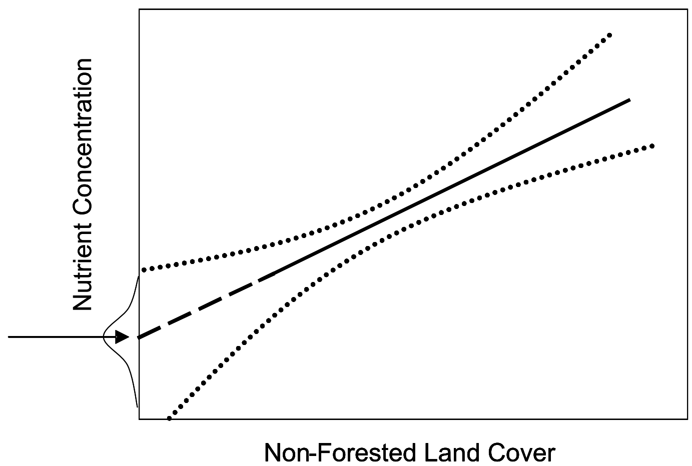
  - Square
  - Cube
  - 10^x

  Negatively skewed

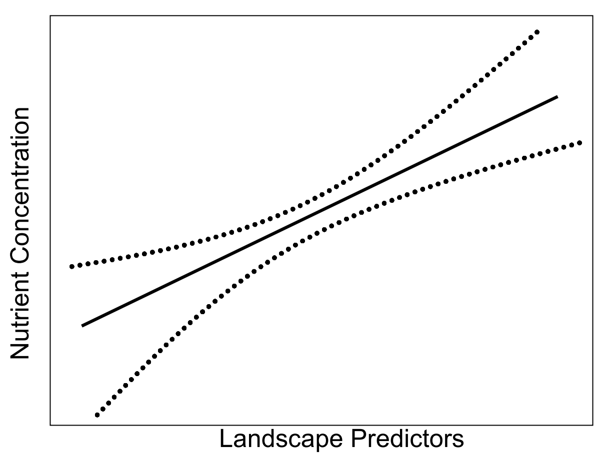
Modeled and Inferred Reference Expectation

- Build models to estimate what the least disturbed nutrient conditions are
- Usually with regression models



TP = Intercept + x(Agriculture) + y(Urban)

Dodds and Oakes 2004



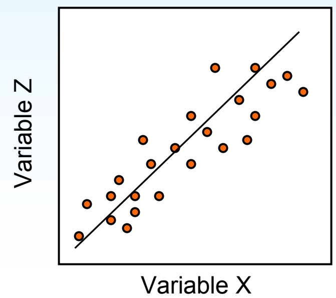
TP = x(lake depth) + y(lake color) + z(lake substrate)

Regression

- Technique that treats one variable as a function of another
- Provides an equation that allows prediction of response from predictor
- Can consider more complex relationships than correlation by using more than two variables

Appropriate questions:

- Can I use chlorophyll a to determine the number of macroinvertebrate taxa in a stream?
- Can surrounding land cover predict the loading of phosphorus to a lake?



Regression – Data Required

- Independent pairs of points
- Numeric variables, may include 0/1 (binary, presence-absence)
- Data where one (or more) variable(s) viewed as predictor(s) and one as response
- Predictors should span as wide a range of values as one might reasonably expect to encounter in the environment

Regression Types

- **Simple linear regression (SLR)**
- Multiple linear regression
- Nonlinear regression
  - Logistic regression
  - Exponential regression
  - Polynomial regression

Linear Regression

- Error term normal with constant variance (check this using the distribution of the response variable)
- Linear relationship (check with plot)
- Model the mean response ($E(Y_i)$)
- Typically use Least Squares estimation
- Variance of each parameter used to evaluate its significance (difference from zero)

$$E(Y_i) = \beta_0 + \beta_1 X_i$$

β_0=intercept
β_1=slope

$$b_1 = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sum (X_i - \overline{X})^2}$$

$$b_0 = \overline{Y} - b_1 \overline{X}$$

Linear Regression Interpretation

- Significance = intercept or slope different from zero
- R^2 = proportion of variability explained
- Examine residuals
- Limit prediction/forecasting

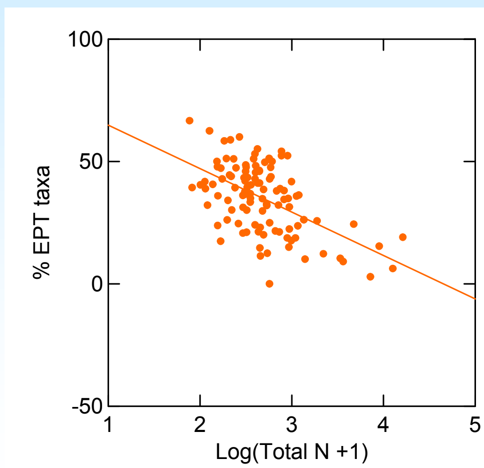
Example – Benthic Response

$F = 44.2$, $p < 0.001$

106 degrees of freedom (df)

R^2=0.287

% EPT = 82.6 – 17.8*Log(TN)



Advantages and Disadvantages

Advantages:

- Predictive
- Interpretation easy
- Lots of types
- Synergies can be modeled with interaction terms

Disadvantages:

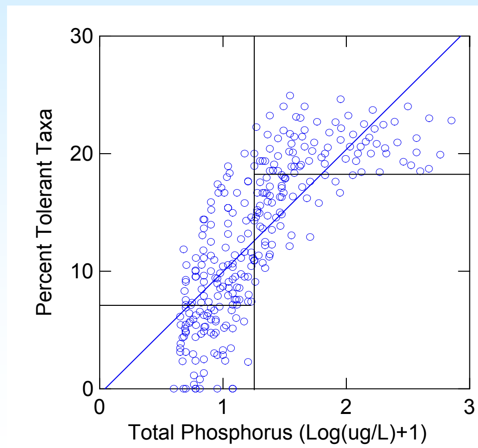
- Assumptions about response variable distribution
- Shape of response used to choose appropriate model
- Interpretation difficult as model becomes more complex (e.g., more predictors, nonlinear)

Regression Alternatives

- Correlation and Regression Trees (CART)
- Generalized Linear Models (GLM)
- Generalized Additive Models (GAM)
- Nonparametric Regression

Stressor-Response

- Many methods
- Choose defensible ones

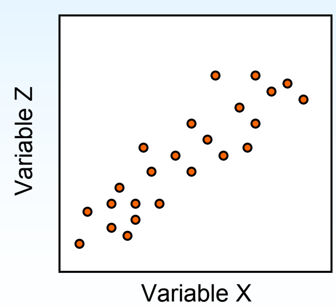


Correlation

- Correlation is a measure of the strength and direction of the relationship between two variables
- Correlation does not indicate causality, but can help to identify variables that covary in the environment

Appropriate questions:

- How strongly is total phosphorus concentration related to the richness of macroinvertebrate taxa?
- Does the amount of agricultural land in the watershed covary with total nitrogen concentration?

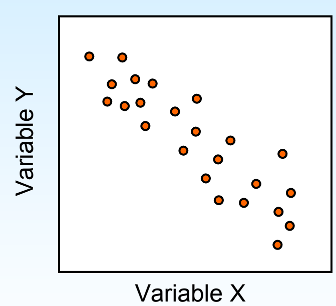


Correlation: Data Required

- Independent data points in the form of paired observations of data (XY pairs)
- Numeric variables required
- Preferably continuous, although discrete numeric variables like richness can work also
- Works best when data spans as much of the range of each variable as possible (or at least close to the range observed in the environment)

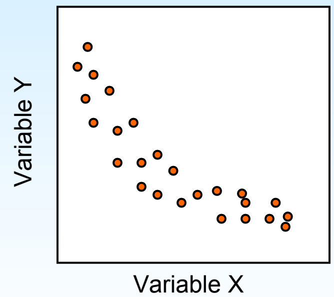
Pearson Correlation

- Assumes approximately normal distribution of both variables and linear relationship
- A linear relationship or normality may be achievable by transformation of one or both variables
- Compares sums of squares (variance) and sums of products (covariance) of the pairs of variables (cov/(varX*varY))



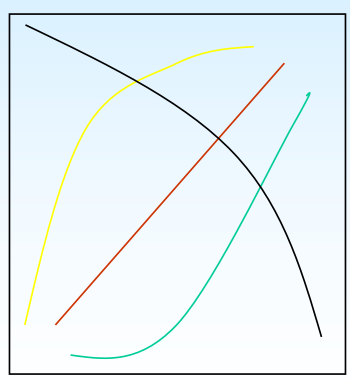
Spearman Correlation

- Just assumes the two variables are random samples from a continuous bivariate population
- Monotonic relationship assumed, can be nonlinear
- Rank values of each variable separately
- Correlation coefficient calculated similarly to that for Pearson but on ranks



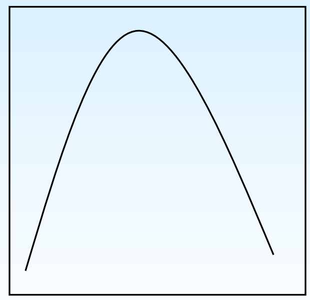
Correlation - Interpretation

- For all measures, range of possible values between -1 and 1
- A significant correlation indicates:
  - A linear association between X and Y
  - The sign of the correlation indicates the direction of the association

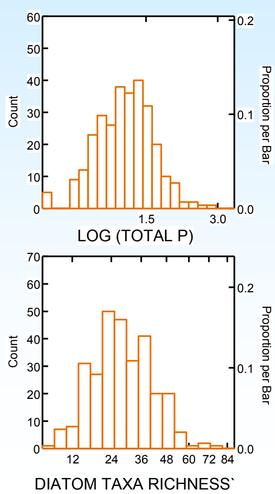


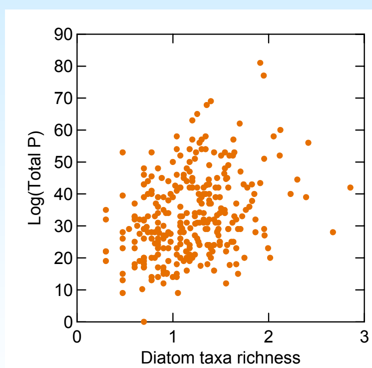
Correlation - Interpretation

- Examination of bivariate plots can be essential for interpretation and in choosing the appropriate kind of correlation to run
- Nonsignificant correlation does not necessarily indicate a lack of association or dependence
  - For Pearson, could occur if relationship nonlinear
  - For Spearman, could occur if relationship non-monotonic



Example – Total P and Diatom taxa richness





- **Pearson r = 0.300, p<0.001, N=294**
- **Spearman r = 0.288**

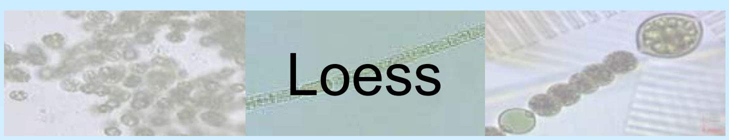
Advantages and Disadvantages

Advantages

- Simple measure that is readily understandable
- Effective to convey simpler monotonic relationships
- Quantitative way to verify bivariate association when used in conjunction with scatter plots

Disadvantages

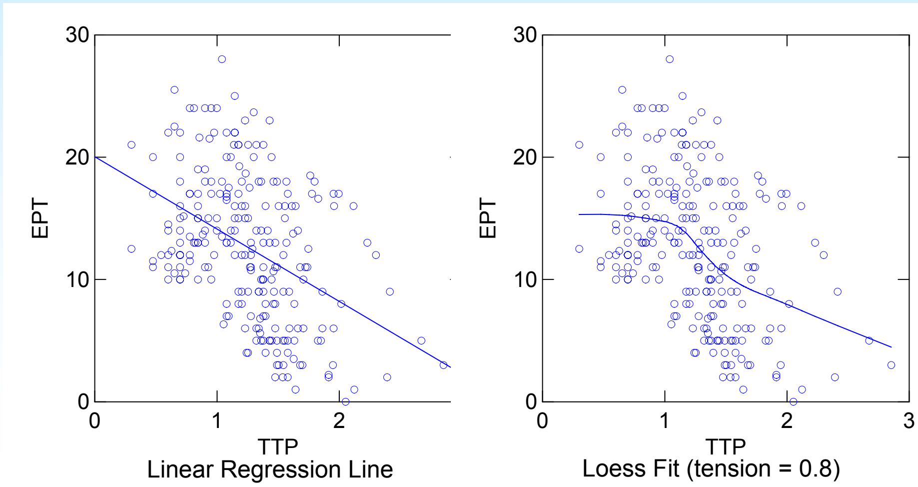
- Cannot detect quadratic or more complex relationships
- Nonsignificant correlation does not necessarily mean no relationship
- Large sample sizes can lead to significant but small correlations
- Not predictive, does not allow forecasting



Loess

- Loess = locally estimated scatterplot smoothing
  - Sometimes seen as Lowess (locally weighted)
- What’s a smoothing function
  - Capture general trends in response
  - Minimal assumptions

Loess

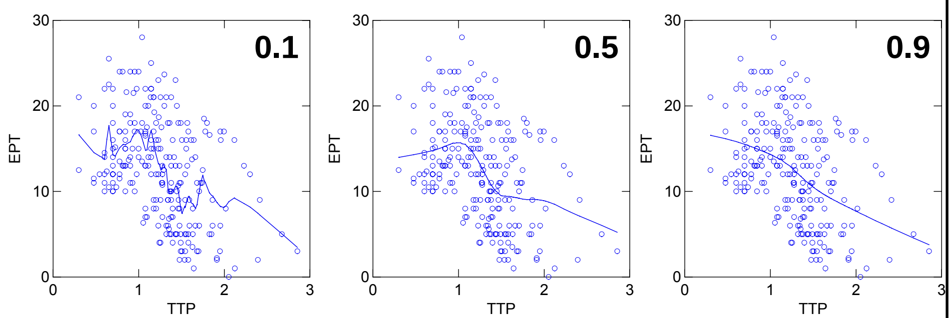


Loess Nuts and Bolts

- Nuts and Bolts
  - Select a set portion of the data (0 to 1)
    - “tension”, “bandwidth”, or “smoothing parameter”
  - Regress a least squares line through those points
    - Linear or simple polynomial
    - Weight the closest points most
    - Use equation to predict value for the first point
  - Shift one point to the right and keep going
  - Connect the points

Loess

- Things you can tweak – bandwidth
- Region of smoothing (0 to 1)



- Can also tweak the type of line fit and weights

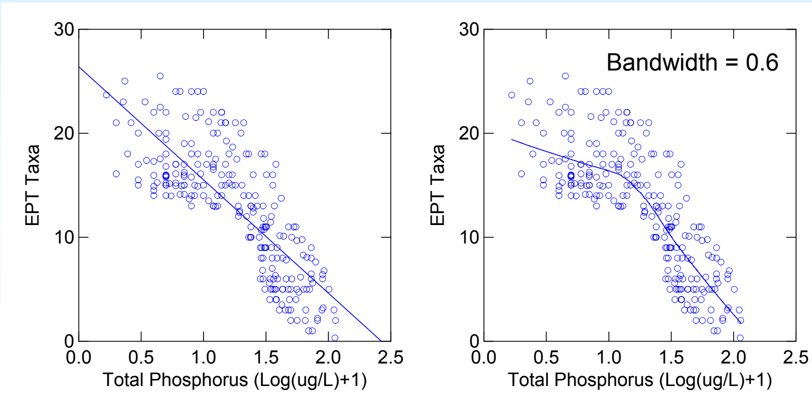
Loess – Data Required

- Pretty much any stressor-response data
- Very little restriction
- No assumptions
- Need a lot of data – the more the better

Advantages and Disadvantages

- Pros
  - Simple
  - Very flexible
  - No assumed relationship
  - Great for visualizing complex relationships
  - User can estimate new value to fit and validate the model if needed
- Cons
  - Requires densely sampled datasets
  - No ready formula produced – hard to transport
  - Computationally intensive (not a problem today)
  - Sensitive to outliers

Loess



Alternatives to Loess

- Least-squares linear regression
- Least-squares non-linear regression
- Smoothing spline
- Distance weighted least squares
- Others

Change-Point Analysis

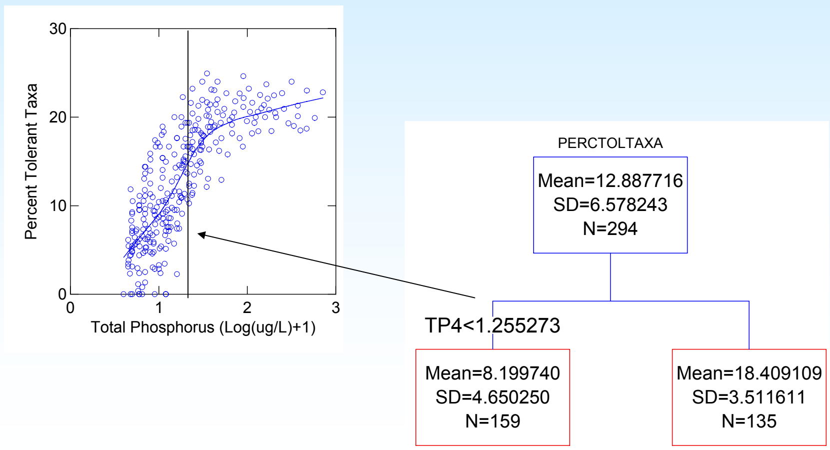
Simply

- Threshold identification method
  - Is there a point where the behavior of values below the point is different than above it?

Less Simply

- A method that attempts to find the point ***r*** along a distribution of values of ***y*** whereby the characteristics of ***y*** (mean and variance) below ***r*** are different than those above ***r***.
  - The value for ***x*** at point ***r*** is the threshold.

Change-Point Analysis



Change-Point Analysis

- Nuts and Bolts
  - A number of methods
    - Non-parametric deviance reduction
      - Find that point where the sum of the deviance on either side is lowest compared to the entire dataset
      - This is a regression tree with <u>one</u> predictor
      - Essentially an iterative method (step by step)
      - Use a bootstrap method to evaluate confidence in threshold

Change-Point Analysis

- Things you can tweak
  - Approach – non-parametric deviance reduction or cumulative sum, bayesian modeling approaches exist too.
  - For non-parametric –
    - Variance method – least squares, deviance reduction, cumulative sums, etc. are the major tweaks
    - How much reduction in error you require

Change-Point Analysis – Data Required

- Continuous or categorical
    - Nutrients vs chlorophyll, metrics, or binomial data – for example – presence/absence of nuisance algae

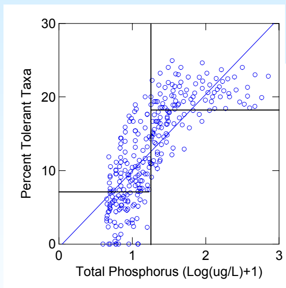
Change-Point Analysis - Reporting

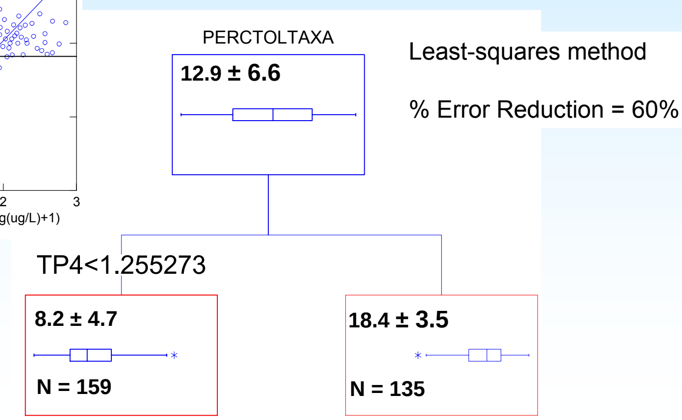
- Important information for interpretation
  - Biggie – error reduction
    - How much error is reduced by selecting that value
    - Some use χ^2 test to evaluate significance of reduction
  - Error around threshold estimate
    - Bayesian techniques produce a range
    - Non-parametrics can produce confidence around estimate (bootstrap methods)

Advantages and Disadvantages

- Pros
  - Identifies clear thresholds
  - Non-parametric is distribution free – few assumptions
  - Can use confidence intervals to set criteria
- Cons
  - Data hungry
  - Computationally intensive
  - Hard with highly variable data
  - Estimate error from re-sampling method (e.g., bootstrap)

Examples – benthic response





Alternatives to Change-Point Analysis

- Not many – threshold analysis is not easy
- Conditional probability analysis

- Visual methods
  - Scatterplots of stressor-response
  - Loess

Conditional Probability

- Probability of observing a condition y for a given value of x

For example:

- What is the probability of observing DO criterion violation for a given concentration of TP?

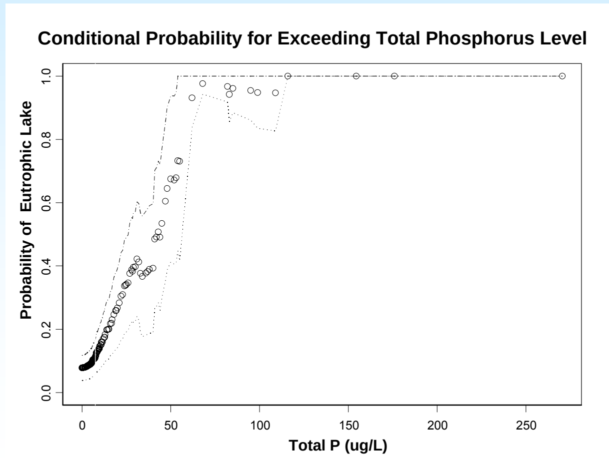
Conditional Probability

P (Y = 1 | X) → probability of observing streams with poor ecological condition (Y = 1) if we observe TP > 20 μg/L (X)

$$P(Y = 1 \mid X) = \frac{\textit{Number of streams in poor condition with turbidity} > 20\ \textit{NTUs}}{\textit{Number of streams with turbidity} > 20\ \textit{NTUs}}$$

Conditional Probability

- **Done over a range of TP concentrations, one can generate a conditional probability curve**
- **Combine with change-point analyses to identify thresholds**



Conditional Probability Analysis

- Nuts and Bolts
  - Calculate P(y = 1| X) over the range of Xs.
  - Convert to conditional cumulative distribution function curve
  - For threshold analysis:
    - Deviance reduction
    - Non-overlapping confidence intervals
    - Visual inspection
    - Change in curvature

Conditional Probability Analysis

- Things you can tweak
  - Threshold approach – 4 different ones
  - Still some debate over best

Conditional Probability Analysis – Data Required

- Matched stressor and response data
- Continuous or categorical
  - Ideally the stressor (x-axis) is continuous
  - Response (y-axis) continuous or categorical
    - Converted into binomial (0 or 1)

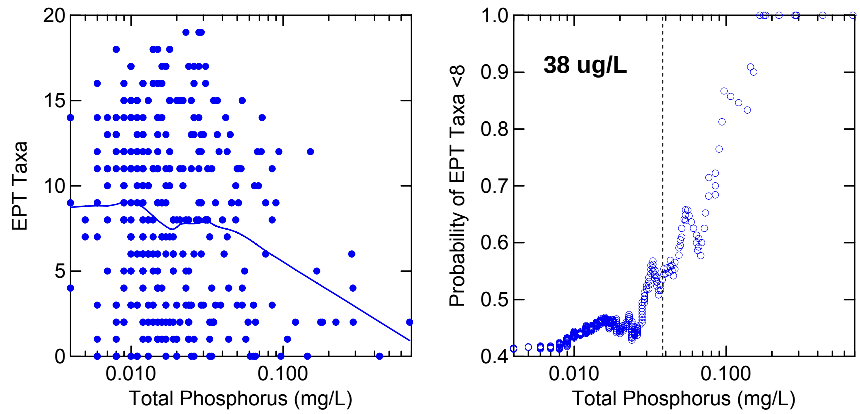
Conditional Probability Analysis - Reporting

- Important information for interpretation
  - Curve
  - Threshold method
  - Threshold estimate
  - Error around threshold estimate (or range)

Advantages and Disadvantages

- Pros
  - Great way to visualize relationship of nutrient to risk
  - Identifies thresholds
  - Fairly distribution free – few assumptions
  - Can use confidence intervals to set criteria
- Cons
  - Preferably probabilistic data
  - Data hungry
  - Hard with highly variable data
  - Estimate error from re-sampling method (e.g., bootstrap)
  - Confounded by other stressors

Conditional Probability Analysis



Models

- Mechanistic
- Empirical

Stream Water Quality Model (QUAL2K)

QUAL2K (or Q2K) is a river and stream water quality model that is intended to represent a modernized version of the QUAL2E (or Q2E) model

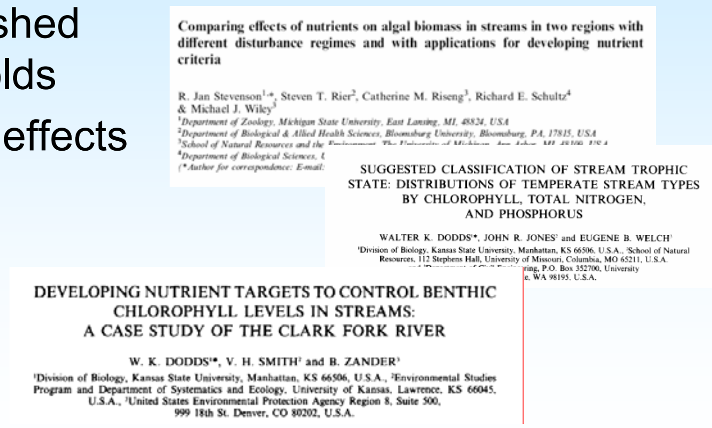
Regional interpretation of water-quality monitoring data

Richard A. Smith, Gregory E. Schwarz, and Richard B. Alexander
U.S. Geological Survey, Reston, Virginia

Abstract. We describe a method for using spatially referenced regressions of contaminant transport on watershed attributes (SPARROW) in regional water-quality assessment. The method is designed to reduce the problems of data interpretation caused by sparse sampling, network bias, and basin heterogeneity. The regression equation relates measured transport rates in streams to spatially referenced descriptors of pollution sources and land-surface and stream-channel characteristics. Regression models of total phosphorus (TP) and total nitrogen (TN) transport are constructed for a region defined as

Other Studies

- Established thresholds
- Known effects levels



Comparing effects of nutrients on algal biomass in streams in two regions with different disturbance regimes and with applications for developing nutrient criteria

R. Jan Stevenson[1,*], Steven T. Rier[2], Catherine M. Riseng[3], Richard E. Schultz[4] & Michael J. Wiley[3]
[1]*Department of Zoology, Michigan State University, East Lansing, MI, 48824, USA*
[2]*Department of Biological & Allied Health Sciences, Bloomsburg University, Bloomsburg, PA, 17815, USA*
[3]*School of Natural Resources and the*
[4]*Department of Biological Sciences,*
(**Author for correspondence: E-mail:*

SUGGESTED CLASSIFICATION OF STREAM TROPHIC STATE: DISTRIBUTIONS OF TEMPERATE STREAM TYPES BY CHLOROPHYLL, TOTAL NITROGEN, AND PHOSPHORUS

WALTER K. DODDS[1]*, JOHN R. JONES[2] and EUGENE B. WELCH[3]
[1]Division of Biology, Kansas State University, Manhattan, KS 66506, U.S.A., [2]School of Natural Resources, 112 Stephens Hall, University of Missouri, Columbia, MO 65211, U.S.A.
ring, P.O. Box 352700, University
e, WA 98195, U.S.A.

DEVELOPING NUTRIENT TARGETS TO CONTROL BENTHIC CHLOROPHYLL LEVELS IN STREAMS: A CASE STUDY OF THE CLARK FORK RIVER

W. K. DODDS[1]*, V. H. SMITH[2] and B. ZANDER[3]
[1]Division of Biology, Kansas State University, Manhattan, KS 66506, U.S.A., [2]Environmental Studies Program and Department of Systematics and Ecology, University of Kansas, Lawrence, KS 66045, U.S.A., [3]United States Environmental Protection Agency Region 8, Suite 500, 999 18th St. Denver, CO 80202, U.S.A.

Approaches

- Multiple Lines of Evidence
  - Frequency distribution approaches
  - Modeled and inferred reference expectation
  - Nutrient-response models
  - Mechanistic/Other Models
  - Other Studies

USEPA 2000

Multiple Lines of Evidence

- Generate candidate endpoints
- Weight qualitatively (BPJ)
- Final is a result of multiple lines

45 µg/L Distribution

40 µg/L Literature

65 µg/L

70 µg/L Stressor-response

55 µg/L Models

Example

Clark Fork River

- Reference Reaches
- Experimental studies – stressor-response
- Scientific Literature
- External Review